What is claimed is:

1. An animation system comprising:

an audio analysis unit having an audio output and a state change output; and

a control unit having an input and an output, the input being constructed to receive a state change indication from the audio analysis unit, the control unit being further constructed to generate compositing command information for individual component elements of an animation frame and send the compositing command information out the output.

2. The animation system of claim 1, further comprising:

a compositing unit constructed to create a composite image sequence, from individual graphic elements, based upon the compositing command information received from the control unit, the composite image sequence being synchronized to an audio stream received from the audio output of the audio analysis unit.

3. The animation system of claim 2, further comprising a cueing unit coupled to the control unit.

4. The animation system of claim 2 wherein the control unit comprises a transition controller and a transition table.

5. The animation system of claim 2 further comprising a graphics database coupled to the compositing unit, the graphics database having graphic information stored for access by the compositing unit.

6. The animation system of claim 5 wherein the graphics database comprises digital images.

7. The animation system of claim 5 wherein the graphics database comprises image definition information.

8. The animation system of claim 2 further comprising a timeline recorder coupled to the control unit.

9. The animation system of claim 2 wherein the control unit further comprises a state change controller.

10. The animation system of claim 9 wherein the state change controller is a trigger controller.

11. The animation system of claim 9 wherein the state change controller is a continuous controller.

12. The animation system of claim 9 wherein the state change controller is a programmed event controller.

13. The animation system of claim 2 wherein the audio analysis unit comprises a signal analysis module.

14. The animation system of claim 13 wherein the signal analysis module is constructed to perform a fast fourier transform on an input audio stream.

15. The animation system of claim 14 further comprising classification bandwidth settings defining between at least two frequency bands and about five frequency bands.

16. The animation system of claim 15 wherein the classification bandwidth settings are tunable.

17. The animation system of claim 14 wherein the signal analysis module is constructed to output spectral content within the at least two frequency bands and about five frequency bands, power and zero crossing information for the input audio stream.

18. The animation system of claim 2 wherein the audio analysis unit comprises a mouth classification unit constructed to receive spectral analysis information and output mouth state information for provision to the control unit.

19. The animation system of claim 2 wherein the audio analysis unit further comprises a classification table.

20. The animation system of claim 19 wherein the classification table associates spectral information with mouth state information.

21. The animation system of claim 2 wherein the audio analysis unit further comprises a signal analysis module and a mouth classification unit.

22. The animation system of claim 21 wherein the signal analysis module is one of at least two signal analysis modules.

23. The animation system of claim 22 wherein the at least two signal analysis modules share the mouth classification unit.

24. The animation system of claim 21 wherein the signal analysis module is one of at least two signal analysis modules and the mouth classification unit is one of at least two mouth classification units.

25. The animation system of claim 24 wherein each signal analysis module is coupled to a mouth classification unit in a one-to-one relationship.

26. The animation system of claim 21 further comprising a multiplexor coupled to the signal analysis module and a demultiplexor coupled to the mouth classification unit so that at least two audio tracks can share the signal analysis module.

27. The animation system of claim 21 wherein the signal analysis module performs time-slice pre-filtering.

28 The animation system of claim 2 wherein the audio analysis unit and compositing unit run on a single computer.

29. The animation system of claim 2 wherein the audio analysis unit runs on of a first computer and the compositing unit runs on a second computer.

30. The animation system of claim 2 wherein each of the audio analysis unit, control unit and compositing unit run on separate computers.

31. The animation system of claim 2 wherein the control unit is constructed to accept MIDI signals.

32. The animation system of claim 2 wherein the audio analysis unit further comprises an audio delay device constructed to output the audio stream.

33. The animation system of claim 2 further comprising:

a processor;

storage, communicatively coupled to the processor; and

a program comprising instructions stored in the storage which, when executing, implements functions of the audio analysis unit.

34. The animation system of claim 2 further comprising:

a processor;

storage, communicatively coupled to the processor; and

a program comprising instructions stored in the storage which, when executing, implements functions of a transition controller.

35. The animation system of claim 2 further comprising:

a processor;

storage, communicatively coupled to the processor; and

a program comprising instructions stored in the storage which, when executing, implements functions of a compositing engine.

36. The animation system of claim 2 further comprising a transition table.

37. The animation system of claim 36 wherein the transition table includes entries to effect an in-pose transition.

38. The animation system of claim 36 wherein the transition table includes entries to effect a pose to pose transition.

39. The animation system of claim 36 wherein the transition table further comprises at least one of a command, a program segment, data, control information, a direct reference to an image, an indirect reference to an image, vector description information, a location, a transformation, a tag, a magnification factor, a zoom factor, a function or a parameter.

40. An animation development system comprising:

a library of graphic elements defining steps in a sequence to effect a change from one state to another state and capable of being composited together into animation frames;

a controller constructed to respond to state change information and generate graphic element change commands in response; and

a compositing engine coupled to the library and the controller so that when the graphic element change commands generated by the controller are received, the controller will access the library and generate composite animation frames incorporating selected graphical elements identified by the steps in the sequence based upon the graphic element change commands.

41. The animation development system of claim 40 wherein the library further comprises character component elements.

42. The animation development system of claim 40 wherein the controller is further constructed to respond to audio analysis state change information.

43. The animation development system of claim 42 wherein the state change information is usable to effect a change in a graphic element other than a mouth.

44. The animation development system of claim 42 wherein the state change information is usable to effect a change in a mouth graphic element.

45. The animation development system of claim 40 wherein the library comprises data representing mouth positions during speech.

46. The animation development system of claim 45 further comprising a delay having an input constructed to accept an audio stream.

47. The animation development system of claim 46 wherein the controller is constructed to respond to audio analysis state change information generated from the audio stream and the delay is constructed to delay the audio stream by an amount sufficient to allow for synchronization of the audio stream with incorporated graphical elements.

48. A method of creating an image stream comprising:

stepping through individual components of a first transition sequence for a first image component element based upon first state change information;

correlating the individual components of the first transition sequence to first stored graphical information representing the first image component element as the first transition sequence is stepped through;

stepping through individual components of a second transition sequence for a second image component element based upon second state change information;

correlating the individual components of the second transition sequence to second stored graphical information representing the second image component element as the second transition sequence is stepped through; and

compositing the first stored graphical information and the second stored graphical information into frames such that, as the first and second transition sequences are stepped through, a sequence of composite frames comprising the image stream is created including the first stored graphical information and the second stored graphical information.

49. The method of claim 48 further comprising performing a format conversion on the image stream.

50. The method of claim 48 further comprising displaying the image stream in real time.

51. The method of claim 48 further comprising recording the image stream.

52. The method of claim 51 further comprising laying the image stream to tape.

53. The method of claim 48 further comprising analyzing an input audio stream.

54. The method of claim 53 further comprising synchronizing the input audio stream to the image stream.

55. The method of claim 53, wherein the first stored graphical information comprises mouth shapes, the method further comprising synchronizing the mouth shapes to the input audio stream based upon the analyzing.

56. The method of claim 53 wherein the analyzing comprises performing a Fast Fourier Transform on the audio stream.

57. The method of claim 53 further comprising correlating mouth images to the audio stream based upon a result of the analysis.

58. The method of claim 53 further comprising pre-filtering the audio stream.

59. An animation system comprising:

a character table;

a gesture table;

a sequence table; and

a transition table, the transition table interrelating the character table, gesture table and sequence table such that, when state change information is received for a character represented in the character table, the state change information representing a character gesture, the state change information will cause a sequencing through entries in the transition table to occur and images identified in the sequence table by the entries to be composited into an image stream.

60. A lip synch method comprising:

analyzing an audio stream to obtain a series of spectral patterns for the audio stream, each of the patterns comprising spectral content within from at least two bands to about five bands and a zero crossings count;

generating a mouth state stream based upon the series of spectral patterns; and correlating the mouth state stream to mouth shape images such that the mouth shape images are synchronized with the audio stream.

61. The method of claim 60 wherein the generating comprises comparing spectral content in frequency bands, spectral power and zero crossings with stored values.

62. The method of claim 61 wherein the frequency bands comprise three frequency bands, the method further comprising identifying a mouth state based upon a spectral content in each of the three frequency bands, a spectral power value and a zero crossing value.

63. The method of claim 60 further comprising testing for a V mouth, followed by an L mouth, then a B mouth, then an A mouth, then an M mouth, then an F mouth, then an E mouth, then either a C mouth or a D mouth.

64. The method of claim 60 further comprising:

testing for a V mouth before testing for any of an F mouth, E mouth, B mouth E mouth or A mouth.

65. The method of claim 60 further comprising:

testing for an F mouth before testing for any of a B mouth, A mouth or M mouth.

66. The method of claim 60 further comprising:

testing for one of a C mouth or a D mouth after testing for an F mouth, B mouth or A mouth.

67. The method of claim 61 further comprising, for a set of ranges corresponding to a mouth,

a) determining if the spectral content in a first band falls within a first range;

b) determining if the spectral content in a second band falls within a second range;

c) determining if the spectral content in a third band falls within a third range;

d) determining if the spectral power falls within a power range;

e) determining if a count of zero crossings falls within a count range; and

if all of a)-e) are satisfied for the mouth, identifying the mouth state for sending, otherwise, performing a)-e) for a new mouth, if the mouth was not a default mouth.

68. The method of claim 67 wherein a)-e) are performed in any order for the mouth.

69. The method of claim 61 further comprising sending mouth states corresponding to mouth shapes. the mouth states having been identified using the spectral content in each of the frequency bands, spectral power and zero crossings.

70. The method of claim 60 wherein the bands comprise exactly four bands.

71. The method of claim 60 wherein the bands comprise exactly five bands.

72. The method of claim 60 further comprising testing for a V mouth, followed by an L mouth, then an F mouth, then an E mouth, then a B mouth, then an A mouth, then an M mouth, then either a C mouth or a D mouth.

73. A data structure for use in generating an animated sequence comprising:

a first set of data elements to define a first character transition from a state to an in-pose goal state;

a second set of data elements to define a second character transition from the state to a new pose goal state, at least some of the data elements being associateable with images such that, when the state is a current state and a new state indication is received

a) if the new state indication identifies the in-pose goal state, the first set of data elements will be sequenced through such that the first data elements will cause images for the first character transition to be inserted into the animated sequence; and

b) if the new state indication identifies the new pose goal state, the second set of data elements will be sequenced through such that the second data elements will cause images for the second character transition to be inserted into the animated sequence.

74. The data structure of claim 73 further comprising a third set of data elements to define a third character transition when the new state transition identifies an undefined goal state.

75. A method of constraining external control of a performance of an animation comprising:

defining a universe of explicit state-to-state transitions for the performance, each of the explicit state-to-state transitions being represented by a defined sequence of identifiers;

receiving information indicating that a performer has triggered a transition from a current state to a new state;

determining if the new state is a defined goal state that is reachable by a state-to-state transition within the universe of explicit state-to-state transitions and, if so, compositing an image sequence into the animation based upon a sequence of identifiers defined for the state-to-state transition, otherwise, compositing a default image sequence into the animation.

76. The method of claim 75 further comprising:

compositing a pre-composited image sequence into the animation.

77. The method of claim 76 wherein the pre-composited image sequence is a procedural.

78 A method of synchronizing mouth images to an input audio stream comprising:

analyzing a portion of the input audio stream to obtain a spectral content, a zero crossings count and a power value for the portion;

accessing pattern data, the pattern data correlating mouth identifiers to sets of analysis information such that each mouth identifier that is a usable mouth indicator corresponds to at least one set of ranges including

i) a specific range of spectral content,

ii) a specific range of zero crossings counts, and

iii) a specific range of power values;

comparing the spectral content, the zero crossings count and the power value for the portion with the pattern data to identify a mouth indicator;

outputting the mouth indicator for the portion in synchrony with the portion.

79. The method of claim 78 wherein the comparing further comprises:

examining the spectral content to determine whether it falls within specific ranges of spectral content.

80. The method of claim 79 further comprising:

following the examining the spectral content, examining the corresponding zero crossings count to determine whether it falls within specific ranges of zero crossings counts.

81. The method of claim 80 further comprising:

following the examining the zero crossings count, examining the corresponding power value to determine whether it falls within specific ranges of power values.

82. The method of claim 78 wherein the comparing further comprises:

identifying the mouth indicator based upon a hierarchical examination of the spectral content, followed by a combination of the zero crossings count and the spectral content, followed by the zero crossings, followed by the power.

83. The method of claim 78 wherein, when a specific set of the sets is satisfied by the spectral content, zero crossings count and power values, the method further comprises outputting one mouth indicator corresponding to the specific set.

84. A computer program stored on a computer readable medium, the program comprising:

a program module to perform the functions of an audio analysis unit.

85. The computer program of claim 84 wherein the program module is constructed to programmatically interact with a transition controller.

86. The computer program of claim 84 wherein the program module comprises instructions that will match audio pattern elements, obtained from analysis of an audio input with stored values that correlate audio patterns to mouth states, to identify a mouth state for the audio pattern elements.

87. The computer program of claim 84 wherein the program module comprises instructions that will perform a fast fourier transform calculation and a zero crossing count for input audio information.

88. A lip synch device comprising:

input circuitry;

data correlating patterns to state information usable to associate specific mouth images with the patterns; and

audio signal analysis circuitry, coupled to the input circuitry, constructed to analyze audio information received by the input circuitry to obtain a stream of patterns and output a stream of state information based upon the stream of patterns.

89. A data structure usable to correlate audio with mouth images, the data structure comprising:

an array of data elements correlating sets of ranges with identifiers of mouth states, the array including a first mouth state identifier correlated to a first set of ranges and a second mouth state identifier correlated to a second set of ranges such that when a set of discrete values obtained from an audio input satisfies the first set of ranges, the first mouth state identifier will be identified and when the set of discrete values obtained from the audio input satisfies the second set of ranges, the second mouth state identifier will be identified.

90. The data structure of claim 89 wherein the first set of ranges comprises a first field to define a range of spectral content and a second field to define a range of zero crossing counts.

91. A system for producing a animation video comprising:

means for receiving a cue input and outputting a state change indication representing a transition end state;

means, coupled to the means for receiving, for determining whether a current state corresponds to the transition end state and for issuing a command when the current state does not correspond to the transition end state;

means, coupled to the means for determining, for recording a timeline of transitions;

compositing means for receiving the command and outputting a composite frame containing a graphical image element associated with the command; and

frame buffer means for receiving the composite frame and outputting it to one of a screen or a recording device.

92 A method of performing a character comprising:

receiving data, resulting from manipulations of a performer, representing changes to be made to component elements of the character in a frame; and

generating a sequence of transition compositing commands based upon the manipulations.

93. The method of claim 92 further comprising:

compositing images for the component elements into a sequence of animation frames, based upon the sequence of transition compositing commands, such that the changes called for by the manipulations of the performer are reflected in the sequence of animation frames as changes in the component elements of the character.

94. A method of creating a performable character from existing image assets, the existing image assets comprising an original sequence of composite frames that change from an original state to a new state as the original sequence is stepped through, the method comprising:

deconstructing pieces of at least some of the composite frames to obtain a series of component image elements extracted from the composite frames and representing a change of a component between one state and another state;

creating a database of the series of component image elements, the database being constructed such that any individual component image element can be accessed independent of any other individual component image element from the series; and

integrating the database into a system such that, when state change information calls for a transition in the component from a first state to a second state, at least some of the individual component image elements in the database for the component will be accessed based upon a transition sequence defined in the system for the transition.

95. The method of claim 94 further comprising:

indexing the database.

96. The method of claim 94 wherein the deconstructing comprises:

masking unwanted elements from a sequence of the existing image assets.

97. A method of creating an animation from a common image inventory, the method comprising:

receiving a first input audio stream during a time period;

receiving first state change indications identifying independent transitions of individual component elements between initial states and new states during the time period;

compositing first component image elements and first mouth component image elements into a first sequence of first individual frames, the first individual frames collectively representing the animation, the first component image elements being identified from the common image inventory based upon the first state change indications and the first mouth component image elements being identified from the common image inventory based upon an analysis of the first input audio stream; and

outputting the animation.

98. The method of claim 97 wherein the animation is one of at least two animations, the method further comprising:

receiving a second input audio stream during the time period; and

compositing the first component image elements and second mouth component image elements into a second sequence of second individual frames, the second individual frames collectively representing a second of the at least two animations, the first component image elements being identified from the common image inventory based upon the first state change indications and the second mouth component image elements being identified from the common image inventory based upon an analysis of the second input audio stream.

99. The method of claim 98 wherein the first input audio stream is in a first language and the second input audio stream is in a second language.

100. The method of claim 98 wherein the compositing the first component image elements and first mouth component image elements into the first sequence of first individual frames and the compositing the first component image elements and the second mouth component image elements into the second sequence of second individual frames occurs during the time period.

101. The method of claim 97 further comprising:

recording the first state change indications.

102. The method of claim 101 further comprising:

receiving a second input audio stream during a second time period;

playing back the recorded first state change indications in synchrony with the second input audio stream; and

compositing the first component image elements and second mouth component image elements into a second sequence of second individual frames, the second individual frames collectively representing a second of the at least two animations, the first component image elements being identified from the common image inventory based upon the recorded first state change indications and the second mouth component image elements being identified from the common image inventory based upon an analysis of the second input audio stream.

103. The method of claim 97 wherein the animation is one of at least two animations, the method further comprising:

receiving second state change indications identifying independent transitions of second individual component elements between second initial states and second new states during the time period; and

compositing second component image elements and first mouth component image elements into a second sequence of second individual frames, the second individual frames collectively representing a second of the at least two animations, the second component image elements being identified from the common image inventory based upon the second state change indications and the first mouth component image elements being identified from the common image inventory based upon an analysis of the first input audio stream.

104. The method of claim 97 further comprising:

following the outputting, editing the animation.

105 A method of simultaneously creating two animations, each having a character sharing common character image elements reflecting actions of the character and differing mouth image elements, the differing mouth image elements comprising first mouth image elements corresponding to a first language in the first animation and second mouth image elements corresponding to a second language in a second animation, the method comprising:

receiving a first audio stream in the first language and a second audio stream in a second language;

analyzing the first audio stream to obtain a first mouth state stream;

analyzing the second audio stream to obtain a second mouth state stream;

identifying a series of first mouth image elements using the first mouth state stream;

identifying a series of second mouth image elements using the second mouth state stream;

compositing a series of character image elements and the series of first mouth image elements together to create a first animation; and

compositing the second series of character image elements and the series of second mouth image elements together to create a second animation.

106. A method of performing a live animation comprising:

receiving a series of signals representing animation element changes based upon movements of a motion capture unit;

correlating individual stored image components to the signals based upon a defined sequence of transition identifiers; and

compositing the individual stored image components into frames of an animation, based upon the defined sequence, such that the movements of the motion capture unit are reflected in the animation as a sequence of the individual stored image components.

107. A method of creating an animation from a set of actions comprising:

receiving a set of signals reflecting the set of actions;

outputting a set of compositing commands based upon the set of signals, the set of compositing commands being related to the set of signals by sequences of transition data items; and

compositing first images into a first animation using the set of compositing commands and a first graphic database.

108 The method of claim 107 further comprising:

storing the set of compositing commands.

109. The method of claim 107 further comprising:

compositing second images into a second animation using the set of compositing commands and a second graphic database.

110. The method of claim 107 further comprising:

editing the stored set of compositing commands to create an edited set of compositing commands; and

compositing second images into a second animation using the edited set of compositing commands and the first graphic database.

111. The method of claim 107 further comprising:

editing the stored set of compositing commands to create an edited set of compositing commands; and

compositing second images into a second animation using the edited set of compositing commands and a second graphic database.

112. A method of creating a library of image elements for use in an animation system comprising:

capturing a series of images within a defined area from several angles;

deconstructing elements from the images, the elements sequentially illustrating a change of a component from a first state to a second state; and

storing at least some of the elements for use in a frames database.

113. A production studio in a box comprising:

a computer having a display and a keyboard; and

a compositing engine configured to receive an editable component element stream and component image information from an image element database, the compositing engine being constructed to composite frames based upon the image information, the computer being further configured to generate the editable component element stream based upon component element state transitions indicated by a state stream, the computer being further configured to provide the editable component element stream to the compositing engine for compositing into the frames.

114. The production studio in the box of claim 113 further comprising a transition unit.

115. The production studio in the box of claim 113 further comprising a transition controller, and wherein the transition controller is configured to receive the state stream.

116. The production studio in the box of claim 115 wherein the transition unit further comprises transition tables accessible by the transition controller.

117. The production studio in the box of claim 115 wherein the state stream comprises an audio state stream.

118. The production studio in the box of claim 115 wherein the state stream comprises a non-audio state stream.

119. The production studio in the box of claim 113 further comprising storage configured to store at least one of the state stream or the editable component element stream.

120. The production studio in the box of claim 113 wherein the computer is a laptop computer.

121. The production studio in the box of claim 113 wherein the laptop computer is a PowerBook.

122. The production studio in the box of claim 113 wherein the computer is an Apple Macintosh computer.

123. The production studio in the box of claim 122 wherein the computer is a PowerBook.

124. The production studio in the box of claim 113 wherein the computer is a Silicon Graphics Computer.

125. The production studio in the box of claim 113 wherein the computer includes a Unified Memory Architecture.

126. A transition table for use in producing an animation comprising:

a first set of entries which, when sequenced through, will effect a first change in a component of an animation from an initial state to a goal state; and

a second set of entries which, when sequenced through, will efect a second change in the component from the initial state to a second goal state.

127. The transition table of claim 126 wherein the transition table further comprises at least one of a command, a program segment, data, control information, a direct reference to an image, an indirect reference to an image, vector description information, a location, a transformation, a tag, a magnification factor, a zoom factor, a function or a parameter.

128. The transition table of claim 126 wherein the first set of entries are stored on a computer readable medium.

129. The transition table of claim 126 further comprising:

a third set of entries which, when sequenced through, will effect a default transition.

130. An animateering apparatus comprising:

an input converter having an output generator constructed to generate output patterns, each pattern being a function of a particular input supplied to the input converter;

a compositing engine;

a set of selectable images of a component, each image being combinable into a composite animation frame by the compositing engine; and

an ordered list of entries, each of the entries identifying a step in a sequence that effects a change to the component as function of a generated output of the input converter when the entries are stepped through, an individual entry in the step identifying an image of the component in a specified state to be provided to the compositing engine from the set of images, so that when an input is applied to the input converter, a corresponding series of composite animation frames is created.

131. The apparatus of claim 130 wherein each of the entries includes an index into the set of selectable images.

132. The apparatus of claim 130 wherein each of the entries includes an image parameter.

133. The apparatus of claim 132 wherein the image parameter comprises at least one location, rotation, translation, scale, spatial orientation parameter.

134. The apparatus of claim 133, wherein the parameter consists of a two-dimensional parameter.

135. The apparatus of claim 133, wherein the parameter consists of a three-dimensional parameter.

136. The apparatus of claim 132 wherein the image parameter comprises at least one color, shading, shadow, or transparency parameter.

137. The apparatus of claim 130 further including an input device connected to the input converter, wherein the input device is at least one of a MIDI board, a signal processor, a waldo, a tracking device, a spot detector, a motion sensor, or a computer keyboard.

138. The apparatus of claim 130 wherein the action is one of multiple actions.

139. The apparatus of claim 138 wherein each of the multiple of actions is contained in a single table for a character.

140. The apparatus of claim 138 further comprising an additional set of selectable images of an additional component, wherein the component and the additional component are different from each other.

141. An animateering method comprising:

specifying a component from among multiple components subject to a parametric change; and

defining an action for the component as an ordered list of data items which can be dynamically sequentially stepped through in an order based upon a change in an input value such that, a first change from a first input value to a second input value will cause a first sequential stepping from a first data item to a second data item and cause a corresponding first parametric

change in the component, and a second change from the second input value to the first input value will cause a second sequential stepping from the second data item to the first data item and a corresponding second parametric change in the component.

142. The method of claim 141 wherein the first sequential stepping comprises *m* sequential steps and the second sequential stepping comprises *n* sequential steps.

143. The method of claim 142 wherein *m* is greater than *n*.

144. The method of claim 142 wherein *m* is less than *n*.

145. The method of claim 142 wherein *m* is equal to *n*.